Amendments to the Specification:

Please add the following new paragraphs between paragraphs [0072] and [0073]:

[0072.1] For example, SNIDE provides a method for predicting one or more locations of single nucleotide polymorphisms in a nucleic acid sequence by calculating a variation frequency from a first base to a second base within a group of bases in a dataset of two or more genes, generating a variation predictiveness matrix from the calculated variation frequency, comparing the nucleic acid sequence one or more groups at a time with the variation predictiveness matrix to assign a variation value to the bases in the nucleic acid sequence, identifying the locations of the bases in the nucleic acid sequence where single nucleotide polymorphisms will likely occur based on the assigned variation value, and outputting the identified locations of the single nucleotide polymorphisms. Note that this method can be implemented as a computer program embodied on a computer readable medium in which each step is performed by one or more code segments.

[0072.2] The dataset of two or more genes may include a known mutation dataset, a dataset of known diseases, a dbSNP database, a non-human mutation database, a disease-specific database, a non-human disease database, a HGMD database, a linkage database, a splice variant database, a translocation database, a database of known mutations, etc. Alternatively, the calculated variation frequency can be adjusted for wild type genes, engineered or non-naturally occurring genes, conservative polymorphisms, non-conservative polymorphisms, cDNA stability, predicted DNA structure, predicted RNA structure, predicted protein structure, post-translational modification sequences, protein stability, predicted protein transport, shuffled genes, site-directed mutagenesis genes, methylated sequences, epigenetic variation, etc. These datasets and terms are well known to those skilled in the art.